Journal of General Virology (2014), **95**, 1338–1351 **DOI** 10.1099/vir.0.063917-0

Pigeon circoviruses display patterns of recombination, genomic secondary structure and selection similar to those of beak and feather disease viruses

Tomasz Stenzel,[1] Tomasz Piasecki,[2] Klaudia Chrząstek,[2] Laurel Julian,[3] Brejnev M. Muhire,[4] Michael Golden,[4] Darren P. Martin[4] and Arvind Varsani[3,5,6,7]

Correspondence
Tomasz Stenzel
tomasz.stenzel@uwm.edu.pl
Arvind Varsani
arvind.varsani@canterbury.ac.nz

[1]Department of Avian Diseases, Faculty of Veterinary Medicine, University of Warmia and Mazury, ul. Oczapowskiego 13, 10-719 Olsztyn, Poland

[2]Department of Epizootiology with Clinic of Birds and Exotic Animals, Faculty of Veterinary Medicine, Wrocław University of Environmental and Life Sciences, 50-360 Wrocław, Poland

[3]School of Biological Sciences, University of Canterbury, Private Bag 4800, Christchurch 8140, New Zealand

[4]Institute of Infectious Diseases and Molecular Medicine, Computational Biology Group, University of Cape Town, South Africa

[5]Electron Microscope Unit, Division of Medical Biochemistry, Department of Clinical Laboratory Sciences, University of Cape Town, Observatory 7700, South Africa

[6]Department of Plant Pathology and Emerging Pathogens Institute, University of Florida, Gainesville, FL 32611, USA

[7]Biomolecular Interaction Centre, University of Canterbury, Christchurch 8140, New Zealand

Pigeon circovirus (PiCV) has a ~2 kb genome circular ssDNA genome. All but one of the known PiCV isolates have been found infecting pigeons in various parts of the world. In this study, we screened 324 swab and tissue samples from Polish pigeons and recovered 30 complete genomes, 16 of which came from birds displaying no obvious pathology. Together with 17 other publicly available PiCV complete genomes sampled throughout the Northern Hemisphere and Australia, we find that PiCV displays a similar degree of genetic diversity to that of the related psittacine-infecting circovirus species, beak and feather disease virus (BFDV). We show that, as is the case with its pathology and epidemiology, PiCV also displays patterns of recombination, genomic secondary structure and natural selection that are generally very similar to those of BFDV. It is likely that breeding facilities play a significant role in the emergence of new recombinant PiCV variants and given that ~50 % of the domestic pigeon population is infected subclinically, all pigeon breeding stocks should be screened routinely for this virus.

Received 18 January 2014
Accepted 13 March 2014

INTRODUCTION

Although pigeon circovirus [PiCV, also referred to as columbid circovirus (CoCV)] was only identified as the likely causal agent of young pigeon disease syndrome (YPDS; also called young bird sickness) in the early 1990s (Woods *et al.*, 1993), it has potentially been causing wasting diseases in young birds since the domestication of pigeons (Woods *et al.*, 1993, 1994). Over the past 20 years, PiCV infections have been documented in many different parts of the world in racing, ornamental and meat pigeons of all ages (Abadie *et al.*, 2001; Coletti *et al.*, 2000; Cságola *et al.*, 2012; Duchatel *et al.*, 2006; Paré *et al.*, 1999; Raue *et al.*, 2005; Stenzel *et al.*, 2012; Woods *et al.*, 1993, 1994).

Many PiCV infections are asymptomatic (Duchatel *et al.*, 2006; Stenzel *et al.*, 2012) with YPDS only occurring in birds up to the age of 8 months (Raue *et al.*, 2005). The classical symptoms of YPDS include lethargy, weight loss, anorexia, regurgitation from the crop, polydipsia and

The GenBank/EMBL/DDBJ accession numbers for the pigeon circovirus genomes determined in this study are KF738843–KF738872.

One supplementary table and two supplementary figures are available with the online version of this paper.

diarrhoea. Sometimes the course of the disease is mild with birds showing no specific symptoms other than poor racing performance, weight loss or respiratory difficulties. Nevertheless, due to the high prevalence of PiCV in both domesticated and feral urban pigeon populations, PiCV infections are one of the biggest health-related issues encountered during the breeding of these birds (Krapez *et al.*, 2012; Raue *et al.*, 2005; Stenzel *et al.*, 2012; Zhang *et al.*, 2011).

Along with the closely related but far better studied species, beak and feather disease virus (BFDV), PiCV is classified within the family *Circoviridae*. Similar to other circoviruses, PiCV has a small circular ssDNA genome ~2000 nt in length that is encapsidated within ~14–17 nm, non-enveloped, icosahedral virions (Biagini *et al.*, 2012). Its two main genes are bidirectionally transcribed from a double-stranded replicative intermediate, with the virion sense gene (*rep*, ORF V1) encoding a replication-associated protein (Rep) and the complementary sense gene (*cp*, ORF C1) encoding a capsid protein (CP) (Mankertz *et al.*, 2000; Todd *et al.*, 2001, 2008).

PiCV genome sequences display ~64 % similarity to those of BFDV and, besides the fact that PiCV infections only very rarely cause feather abnormalities (Raue *et al.*, 2005; Woods & Latimer, 2000), the diseases caused by these viruses are similar in many key respects. In common with BFDV and other known circoviruses, PiCV infections appear to result in a degree of immunosuppression which leaves infected individuals susceptible to viral (e.g. pigeon herpesvirus) and/or bacterial (e.g. *Escherichia coli* or *Klebsiella pneumoniae*) secondary infections (Abadie *et al.*, 2001; Raue *et al.*, 2005; Stenzel *et al.*, 2012; Woods *et al.*, 1993, 1994). As with BFDV, PiCV is transmitted mainly horizontally through the ingestion or inhalation of virus-contaminated faecal material and feather dust (Woods *et al.*, 1993). Also in common with BFDV is the possibility that PiCV might be transmissible vertically from females to their embryos (Duchatel *et al.*, 2005, 2006; Rahaus *et al.*, 2008). PiCV has also been detected in the testes of adult pigeons, suggesting that cocks might be capable of sexually transmitting the virus to females (Duchatel *et al.*, 2006).

Whereas the global dissemination of BFDV has likely been driven in recent times by the international trade in parrots (Harkins *et al.*, 2014; Heath *et al.*, 2004; Julian *et al.*, 2012, 2013; Massaro *et al.*, 2012), the long-distance movement of PiCV is possibly facilitated by both the international pigeon trade and pigeon racing competitions. During preparations for pigeon races, the practice of communally transporting, feeding and watering birds from disparate lofts might create particularly ideal conditions for the transmission and long-distance dissemination of PiCV. Such practices could, by bringing together divergent PiCV lineages and facilitating mixed infections of these, also encourage the recombinational generation of novel PiCV variants. Although it is apparent that recombination is particularly common amongst BFDV isolates found infecting birds within psittacine breeding facilities (Julian *et al.*, 2013), it is presently unknown to what extent recombination occurs amongst PiCV isolates that infect wild and captive pigeons.

Given the many similarities that exist between BFDV and PiCV, we thought that it would be interesting to directly compare genome-wide patterns of diversity, recombination and natural selection that occur in these species. Towards this end we isolated and sequenced a diverse array of 30 new PiCV genomes from Polish pigeons and analysed these together with the full genomes of 17 PiCV isolates and 204 BFDV isolates obtained from public sequence databases. We show that, consistent with their pathological and epidemiological similarities, PiCV and BFDV display strikingly similar patterns of molecular evolution.

RESULTS AND DISCUSSION

PiCV infection frequencies

Of the 324 pigeons examined, 64 % of domestic pigeons and 44.7 % of feral pigeons were determined to be positive for PiCV infection using primers targeting the ~325 bp region of the *cp* gene. This PiCV incidence is similar to that found by Cságola *et al.* (2012) amongst 116 racing pigeons (57 % PiCV-positive) screened in Hungary. Almost half (53 %) of the pigeons that were PiCV-positive had no notable disease symptoms. Pigeon herpesvirus and *Salmonella typhimurium* was detected in four cases, whereas *Chlamydia psittaci* was detected in five cases (Table S1, available in the online Supplementary Material). Methods and results of bacteriological and parasitological screening for all PiCV-positive birds are summarized in Table S1.

PiCV full genome sequence diversity and phylogenetic analysis

Using abutting primer pairs we collectively recovered 30 complete PiCV genomes from carrier pigeons (n=8), fancy pigeons (n=12) and feral pigeons (n=10); see Table 1 for details. These 30 Polish PiCV genomes all displayed >86 % genome-wide pairwise similarity to one another and >83 % similarity to the 17 other complete PiCV genomes available in GenBank which have been sampled from various other parts of the world, including Europe, North America, Africa and Australia (Fig. 1). This degree of species-wide diversity is similar to that observed for BFDV (Julian *et al.*, 2013). Also similar to BFDV, the Rep sequences of these sampled PiCV isolates are more conserved than their CP sequences (see Figs S1 and S2; Varsani *et al.*, 2011; Julian *et al.*, 2013).

A neighbour-joining phylogenetic tree depicting the possible evolutionary relationships between the 47 PiCV genomes indicated that the isolates fall into five major clusters (Fig. 2a) with the largest cluster containing ~57 % of the PiCV genomes, and including viruses sampled from Europe (n=24), North America (n=2) and Africa (n=1). Cságola

Table 1. GenBank accession numbers, isolate names, hosts and country of origin of PiCV genomes recovered in this study (prefix PL) and those available in public databases

Isolate	Sampling year	Accession no.	Country	Host	Breed name	Origin of parent stock
PL44A	2003	KF738857	Poland (Dolnośląskie)	Carrier pigeon		
PL48	2011	KF738859	Poland (Warmińsko-Mazurskie)	Carrier pigeon		Poland
PL53	2011	KF738860	Poland (Warmińsko-Mazurskie)	Carrier pigeon		Poland
PL94	2011	KF738872	Poland (Warmińsko-Mazurskie)	Carrier pigeon	Janssen	Poland, Germany, Belgium
PL172	2012	KF738849	Poland (Warmińsko-Mazurskie)	Carrier pigeon		
PL177	2012	KF738850	Poland (Mazowieckie)	Carrier pigeon		Poland, Germany, Belgium
PL170	2012	KF738848	Poland (Warmińsko-Mazurskie)	Carrier pigeon		Germany, Poland
PL7	2003	KF738869	Poland (Dolnośląskie)	Carrier pigeon		
PL102	2011	KF738843	Poland (Mazowieckie)	Fancy pigeon	Danziger highflier	Poland, Germany
PL201	2013	KF738854	Poland (Pomorskie)	Fancy pigeon	Felegyhazer	
PL197	2013	KF738853	Poland (Mazowieckie)	Fancy pigeon	Danziger highflier	Poland, Germany
PL114	2011	KF738844	Poland (Mazowieckie)	Fancy pigeon	Polish longfaced flyer	Poland
PL66A	2011	KF738866	Poland (Pomorskie)	Fancy pigeon	Danziger highflier	Poland
PL89X	2011	KF738871	Poland (Mazowieckie)	Fancy pigeon		Poland
PL124	2011	KF738845	Poland (Mazowieckie)	Fancy pigeon	Polish longfaced flyer	Poland
PL188	2012	KF738851	Poland (Warmińsko-Mazurskie)	Fancy pigeon	Fantail	Germany, USA, Poland
PL189	2012	KF738852	Poland (Warmińsko-Mazurskie)	Fancy pigeon	Fantail	Germany, USA, Poland
PL57	2011	KF738861	Poland (Warmińsko-Mazurskie)	Fancy pigeon		Poland
PL67	2011	KF738868	Poland (Warmińsko-Mazurskie)	Fancy pigeon	Fantail	Poland, Germany, USA
PL89	2011	KF738870	Poland (Mazowieckie)	Fancy pigeon		Poland
PL63	2002	KF738865	Poland (Dolnośląskie)	Feral pigeon		
PL14	2002	KF738847	Poland (Dolnośląskie)	Feral pigeon		
PL43	2002	KF738856	Poland (Dolnośląskie)	Feral pigeon		
PL13	2002	KF738846	Poland (Dolnośląskie)	Feral pigeon		
PL40	2002	KF738855	Poland (Dolnośląskie)	Feral pigeon		
PL44B	2002	KF738858	Poland (Dolnośląskie)	Feral pigeon		
PL58	2002	KF738862	Poland (Dolnośląskie)	Feral pigeon		
PL60	2002	KF738863	Poland (Dolnośląskie)	Feral pigeon		
PL62	2002	KF738864	Poland (Dolnośląskie)	Feral pigeon		
PL66B	2002	KF738867	Poland (Dolnośląskie)	Feral pigeon		
SRK/US/01	2003	EU840176	USA	Feral pigeon		
Bel 936	Unknown	DQ915956	Belgium	Pigeon		
98-324	1998	JX901125	Belgium	Pigeon		
Bel 18	Unknown	DQ915957	Belgium	Pigeon		
Bel 20	Unknown	DQ915958	Belgium	Pigeon		
zj2	Unknown	DQ090944	China	Pigeon		
zj1	Unknown	DQ090945	China	Pigeon		
fj1	2009	JN183455	China	Pigeon		
Fra A40042	Unknown	DQ915960	France	Pigeon		
De	2000	AF252610	Germany	Pigeon		
Ita 4B	Unknown	DQ915950	Italy	Pigeon		
9030	Unknown	AJ298229	North Ireland	Pigeon		
7050	Unknown	AJ298230	North Ireland	Pigeon		
US 93A	Unknown	DQ915961	USA	Pigeon		
US 002180	Unknown	DQ915962	USA	Pigeon		
Dove	Unknown	DQ915959	Australia	Senegal dove		
NGchicken38	2009	HQ738642	Nigeria	Chicken		

et al. (2012) identified five groups based on the *cp* sequences, which are similar to those we have noted here. Although there was some evidence of sequences clustering within the phylogenetic tree based on their geographical origins (note the large clade of closely related Polish sequences indicated by an arrow in Fig. 2a), this pattern might simply reflect an obvious European sampling bias (87 % of the analysed isolates were sampled in Europe). There was also no

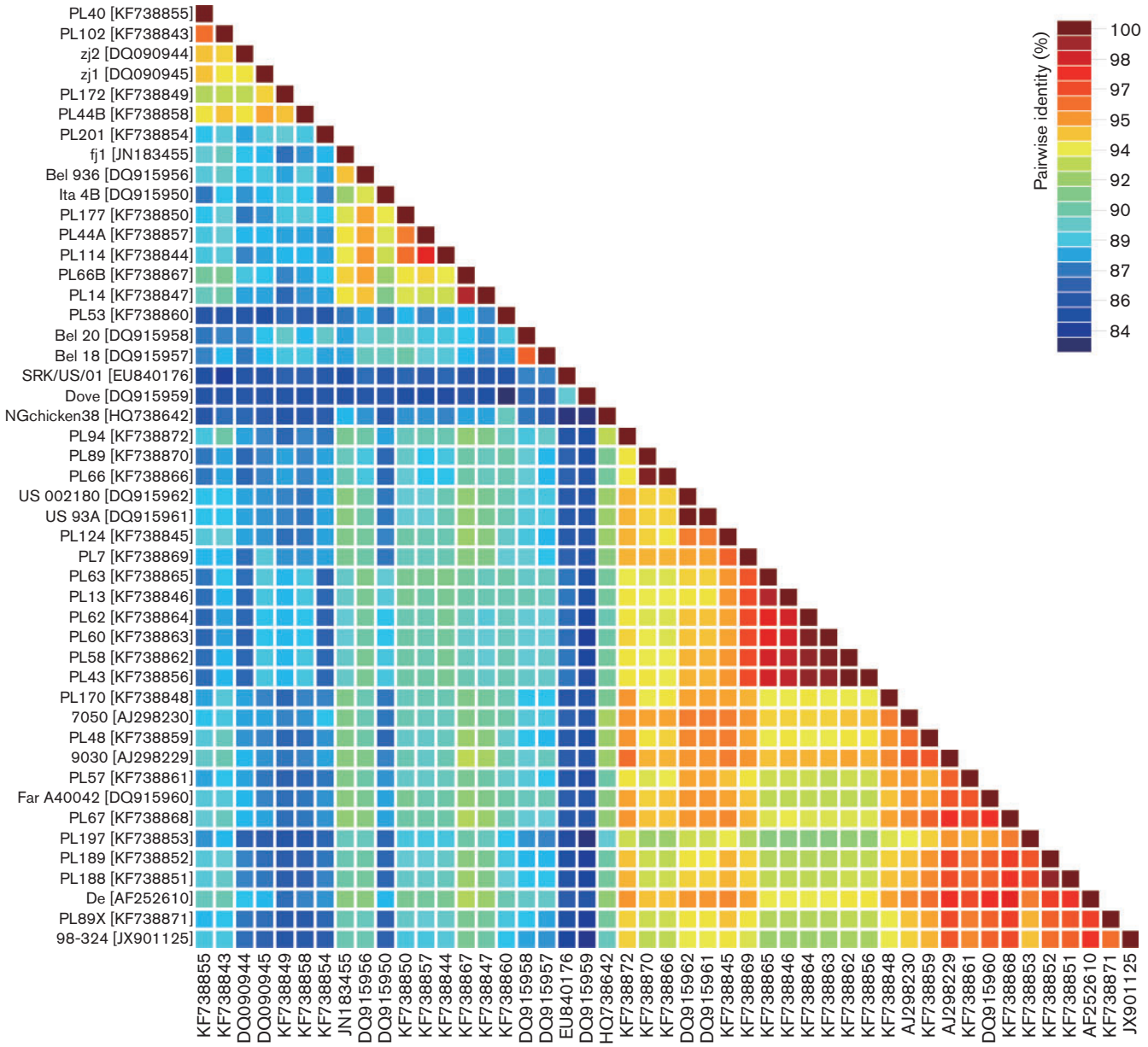

Fig. 1. Two-dimensional genome-wide pairwise identity plot of 47 PiCV isolates calculated using SDT v1.0 (Muhire *et al.*, 2013).

evidence of clustering within the phylogenetic tree of sequences based on the host genotypes (carrier, fancy or wild pigeons) from which they were sampled – a result corroborating that found by Todd *et al.* (2008). This indicates that as with BFDV, long-distance geographical movements and movements between different host genotypes are likely to have occurred relatively frequently during the evolution of PiCV.

Recombination analysis

In general, recombination has played a significant role in the evolution of many ssDNA viruses and has been well documented for certain circoviruses (Cheung, 2009; Julian *et al.*, 2013; Lefeuvre *et al.*, 2009; Varsani *et al.*, 2011). Accordingly, we detected 23 unique recombination events (Table 2; Fig. 2b) within the analysed PiCV genomes, and in some cases detected evidence of two independent events within individual genomes. In the majority of genomes recombinant tracts between 421 and 985 nt were detected. However, in six PiCV genomes (two from North America, one from Australia and three from Europe) we were unable to detect any evidence of recombination based on currently analysed sequences.

A maximum-likelihood phylogenetic analysis of the PiCV genomes with recombinant regions removed indicated that

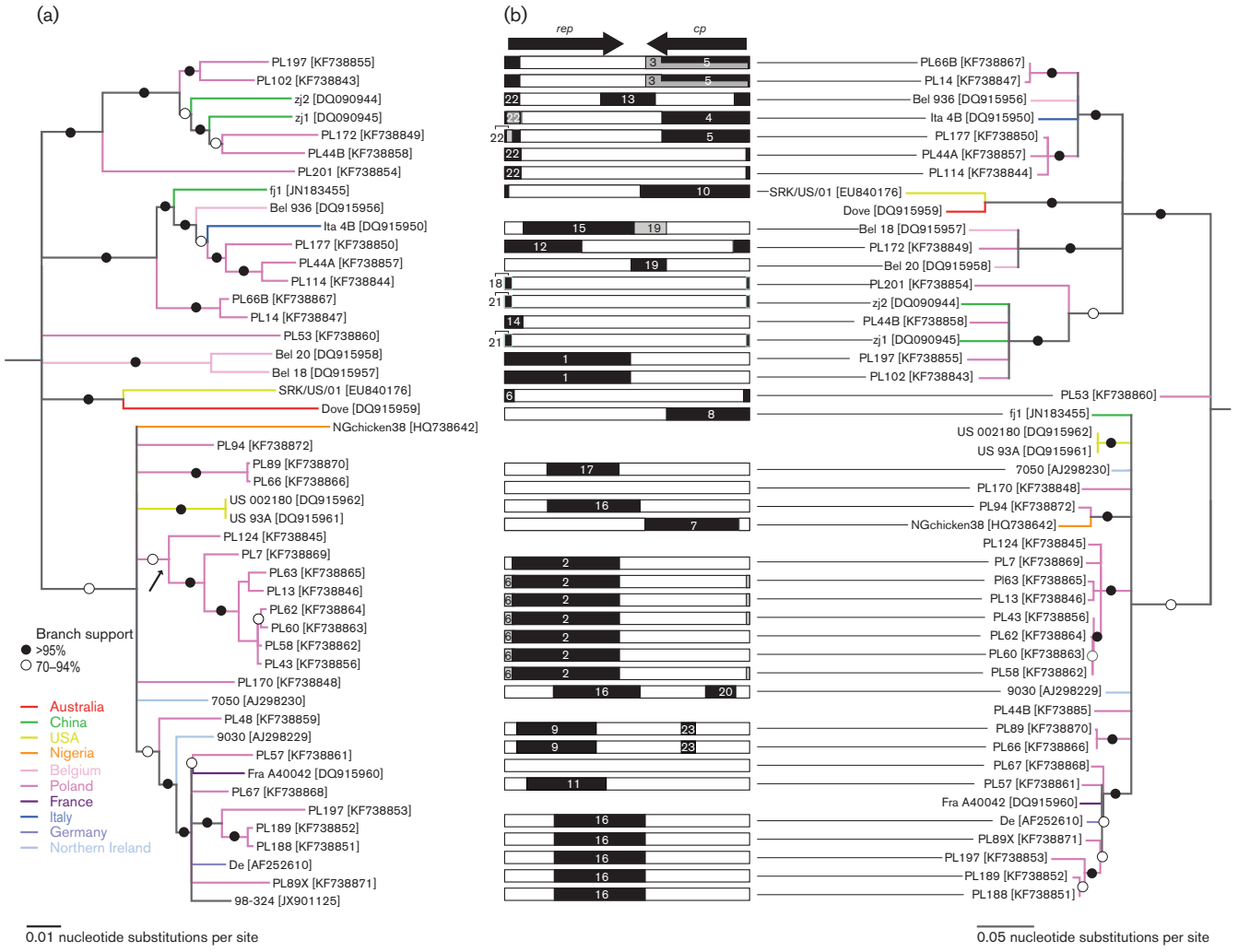

Fig. 2. (a) Neighbour-joining phylogenetic tree illustrating the possible evolutionary relationships between currently sampled PiCV full genome sequences. (b) Cartoon illustration of recombination events detected in PiCV genomes and a maximum-likelihood phylogenetic tree of PiCV genomes with recombinant regions removed. Arrows indicate some evidence of sequence clustering of Polish PiCV isolates; however, this may be attributed to sampling bias.

there exist at least two main PiCV lineages with a third divergent lineage being represented by a single isolate (PL53) from Poland (Fig. 2b).

Analysis of the recombination breakpoint distribution detected during the recombination analysis indicates that recombination patterns in PiCV closely mirror those seen previously in BFDV. Both species have recombination breakpoint hot-spots at similar genomic locations within both the intergenic region between the *rep* and *cp* stop codons, and near the virion strand origin of replication (Fig. 3). They also both have recombination breakpoint cold-spots within the central region of *cp* – a fact that might explain why relatively little evidence for recombination was identified previously in a PiCV diversity study that focused exclusively on the *cp* region (Cságola *et al.*, 2012).

Conservation of genomic secondary structure between PiCV and BFDV

In addition to the well-known structural element at the origin of replication (*ori*), here detected as conserved significantly in both PiCV and BFDV genomes (median *P* value=0.035; Fig. 4), we identified two additional genomic secondary structural elements that are apparently conserved highly in both species: one near the 5′ end of *rep* (median *P* value=0.038), and the other within the intergenic region between the *rep* and *cp* stop codons (median *P* value=0.016). These three potential structural elements had been identified previously and were ranked among the top six out of 35 PiCV and 41 BFDV structural elements detected to be evolutionarily conserved in each species (Muhire *et al.*, 2014). These rankings accounted for base-pairing conservation scores, synonymous substitution

Table 2. Details of recombination events (GenBank accession numbers) detected in PiCV isolates

Methods used to detect recombination are RDP (R) GENCONV (G), BOOTSCAN (B), MAXCHI (M), CHIMERA (C), SISCAN (S) and 3SEQ (T). The method with the most significant associated *P* value is indicated in bold for each event.

Event	Recombinant		Region	Potential minor parent(s)				Potential major parent(s)				Detection methods	*P* value
1	KF738855		31–1009	KF738847	DQ915961	KF738852	AF252610	KF738858				RGMC**T**	8.97×10^{-33}
	KF738843			KF738867	DQ915962	KF738853	AJ298229	DQ090945					
				KF738872	KF738845	KF738859	JN183455	KF738849					
				DQ915960	KF738851	KF738868	KF738871						
2	KF738869	KF738863	95–1080	DQ090945				KF738847	DQ915960	KF738852	AJ298229	RGMC**T**	1.56×10^{-19}
	KF738846	KF738864						KF738867	DQ915961	KF738853	AJ298230		
	KF738856	KF738865						KF738872	DQ915962	KF738859	KF738848		
	KF738862							AF252610	KF738845	KF738861	KF738851		
								KF738871	KF738868				
3	KF738867		1158–2008	KF738850	KF738844			AF252610	KF738845	KF738853	KF738871	RGMC**T**	7.49×10^{-20}
	KF738847			DQ915950				AJ298229	KF738848	KF738859	DQ915960		
				DQ915956				AJ298230	KF738851	KF738861	DQ915961		
				KF738857				KF738872	KF738852	KF738868	DQ915962		
4	DQ915950		1321–140*	KF738844				DQ915959	KF738851	KF738861	DQ915960	RGMC**T**	3.30×10^{-15}
				DQ915956				KF738872	KF738852	KF738866	KF738848		
				KF738857				AF252610	KF738853	KF738868	KF738870		
								AJ298230	KF738859	KF738871			
5	KF738850		1318–132	KF738844				KF738872	DQ915962	KF738861	DQ915960	R**M**CT	1.20×10^{-14}
	KF738847			KF738857				AF252610	KF738848	KF738866	DQ915961		
	KF738867							AJ298229	KF738851	KF738868	KF738853		
								AJ298230	KF738852	KF738870	KF738859		
								KF738871					
6	KF738860	KF738863	2008–85*	KF738849				KF738872	DQ915962	KF738861	DQ915960	RGMC**T**	1.52×10^{-16}
	KF738846	KF738864		DQ915957				AF252610	KF738848	KF738866	DQ915961		
	KF738856	KF738865		DQ915958				AJ298229	KF738851	KF738868	KF738853		
	KF738862			EU840176				AJ298230	KF738852	KF738869	KF738859		
								KF738871	KF738870				
7	HQ738642		1175–1956*	KF738872	DQ915962	KF738859	DQ915960	Unknown				RGM**T**	4.37×10^{-17}
				AF252610	KF738848	KF738861	DQ915961						
				AJ298229	KF738851	KF738868	KF738853						
				AJ298230	KF738852	KF738870	KF738871						
8	JN183455		1351*–8	KF738871	KF738859	DQ915961	AF252610	DQ915950				GMC**T**	9.68×10^{-15}
				KF738870	KF738853	DQ915960	KF738872	DQ915956					
				KF738869	KF738852	AJ298230	KF738848	KF738850					
				KF738868	KF738851	AJ298229	DQ915962						
				KF738866	KF738861								
9	KF738870		89*–752	DQ090945				KF738871	DQ915962	AJ298229	KF738853	R**M**CT	2.29×10^{-11}

Table 2. cont.

Event	Recombinant		Region	Potential minor parent(s)				Potential major parent(s)				Detection methods	*P* value
	KF738866			KF738858				KF738859	DQ915960	AF252610			
10	EU840176		1142–35	DQ915959				DQ090944	KF738858			RMC**T**	4.88×10^{-17}
								DQ090945	DQ915958				
								KF738849					
11	KF738861		178–849	KF738848	KF738853			KF738870				**M**CT	2.37×10^{-6}
12	KF738849		1915–646	DQ915958				KF738854	DQ915962	KF738861	AJ298230	R**M**CT	1.19×10^{-9}
				DQ915957				KF738872	KF738848	KF738868	DQ915960		
								AF252610	KF738852	KF738869	KF738859		
								AJ298229	KF738853	KF738870	KF738871		
13	DQ915956		805*–1266	KF738860	KF738846	KF738863	DQ915960	Unknown				R**G**MCT	2.57×10^{-11}
				KF738872	KF738848	KF738864	DQ915961						
				AF252610	KF738851	KF738865	DQ915962						
				AJ298229	KF738852	KF738867	KF738845						
				AJ298230	KF738855	KF738868	KF738872						
				KF738861	KF738858	KF738869							
				KF738862	KF738860	KF738870							
14	KF738858		9*–156*	DQ915958	DQ915957			DQ090945	DQ090944			**R**GMCT	9.99×10^{-10}
15	DQ915957		157–1092	KF738872	DQ915961	DQ915960	DQ915962	Unknown				RGMC**T**	1.10×10^{-12}
16	KF738872	KF738852	369*–1142	Unknown				AJ298230				RG**M**CT	1.49×10^{-8}
	AF252610	KF738853						DQ915961					
	AJ298229	KF738871						DQ915962					
	KF738851							KF738848					
17	AJ298230		352–962*	EU840176	KF738844			KF738843	KF738848	KF738868		RGMC**T**	2.07×10^{-9}
				DQ915956	KF738857			KF738855	KF738859				
18	KF738854		2014–87	KF738872	DQ915962	KF738853	DQ915960	Unknown				R**G**MC	2.07×10^{-7}
				AF252610	KF738845	KF738859	DQ915961						
				AJ298229	KF738848	KF738861	KF738852						
				AJ298230	KF738851	KF738868	KF738871						
19	DQ915958		1059–1358	Unknown				KF738844				**R**GM	1.35×10^{-6}
	DQ915957												
20	AJ298229		1676–1932*	KF738853				DQ915961	KF738856	KF738865	KF738846	RGMC**T**	7.36×10^{-8}
								DQ915962	KF738862	KF738869	KF738864		
								KF738845	KF738863				
21	DQ090945		2010*–64	KF738872	KF738848	KF738866	KF738853	DQ915950	KF738850	DQ915957		**R**GM	8.61×10^{-6}
	DQ090944			AJ298230	KF738851	KF738869	KF738859	DQ915956	KF738857	DQ915958			
				DQ915961	KF738852	KF738870	KF738845	KF738844					
				DQ915962									
22	DQ915956		1915–132	DQ915958				KF738872	DQ915962	KF738859	DQ915961	R**G**MC	5.73×10^{-7}
	DQ915950			DQ915957				AF252610	KF738845	KF738866	KF738853		

Table 2. cont.

Event	Recombinant	Region	Potential minor parent(s)			Potential major parent(s)		Detection methods	*P* value
	KF738844					AJ298229	KF738848	KF738868	
	KF738857					AJ298230	KF738851	KF738870	
	KF738850					DQ915960	KF738852	KF738871	
23	KF738866	1465–1590	DQ090944	KF738849	KF738854	KF738843	KF738851	KF738853	RMCT	1.03×10^{-4}
	KF738870		DQ090945	KF738855	DQ915959	KF738858	KF738852			

*The actual breakpoint position is undetermined.

rates at paired sites falling within coding regions and evidence of complementary coevolution favouring maintenance of base pairing.

The *rep* structures in these species have distinct conformational similarities even though the sequences within the structures are moderately diverged. The structure within the intergenic region between the *rep* and *cp* stop codons is a relatively simple stem–loop that has been described previously (Muhire *et al.*, 2014). Whereas the loop sequence of this structure contains a conserved pentanucleotide, CGAAG, the 10–15 nt stem sequence is GC-rich and displays strong evidence that base-paired nucleotides are coevolving with one another in a manner which conserves base pairing (Muhire *et al.*, 2014).

Conserved signals of natural selection between PiCV and BFDV

Although both *rep* and *cp* in both PiCV and BFDV are evolving under selection that predominantly disfavours amino acid substitutions (i.e. negative selection as indicated by gene-wide $dN/dS<1.0$ in Fig. 5), specific codon sites within these genes are detectably evolving under positive selection (i.e. selection favouring amino acid substitutions indicated by sites in red and orange in Fig. 5). For both PiCV and BFDV there are over twice as many sites detectably evolving under positive selection in *cp* (10 and 27, respectively) than there are in *rep* (two and 11, respectively). Adaptive host immunity is one of the primary drivers of positive selection in vertebrate infecting viruses (Bennett *et al.*, 2006; Esteban & Hutchinson, 2011; Rigby *et al.*, 1993; Seibert *et al.*, 1995) and these results are therefore expected given that the CP is likely far more exposed to host immune systems than is Rep (Blanchard *et al.*, 2003; Pogranichnyy *et al.*, 2000). Although it might seem from these results that BFDV has many more codon sites evolving under positive selection than does PiCV, this discrepancy is accounted for by the fact that the BFDV dataset is five times larger than the PiCV dataset and that the BFDV analysis therefore had far greater power to detect sites evolving under positive selection.

Similarities in the selection pressures experienced by PiCV and BFDV are perhaps best reflected by similarities in the distributions of negatively selected codon positions across *rep* and *cp* (indicated in green and blue in Fig. 5). At these sites selection favours a specific expressed amino acid and therefore disfavours non-synonymous substitutions.

Surprisingly, there are a large number of sites (26 in *rep* and 23 in *cp*) at which PiCV and BFDV are evolving under negative selection that favour different amino acids in the different species (sites indicated in green in Fig. 5). This result implies that during the time following the divergence of PiCV and BFDV from their most recent common ancestor, directional positive selection (i.e. selection favouring particular amino acids at particular sites in these proteins) has likely been acting at these sites. Amongst these directionally selected sites are likely to be those

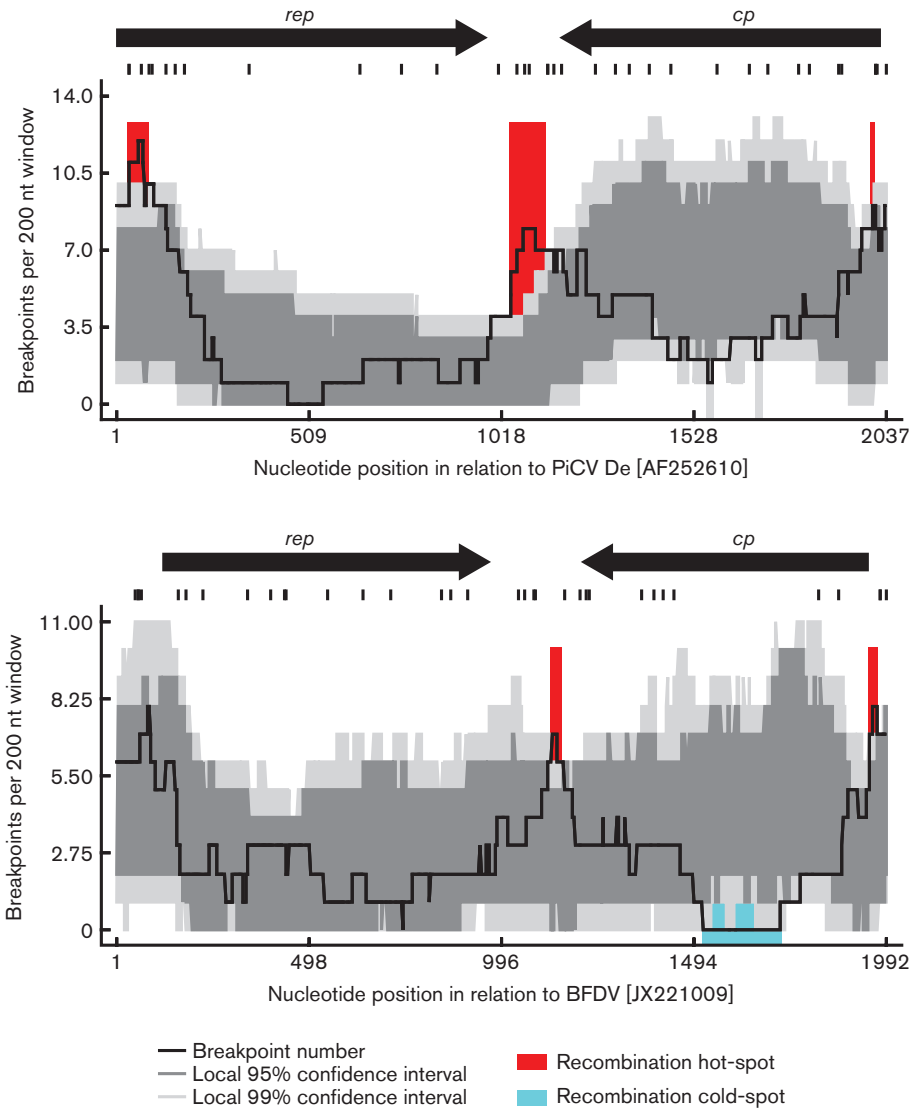

Fig. 3. Recombination breakpoint distribution plots for PiCV and BFDV. Recombination breakpoint hot-spots are indicated in red and cold-spots are indicated in blue. The dark and light grey areas of the plots indicate 95 and 99 % confidence intervals, respectively, on the expected density of detectable recombination breakpoints under random recombination in the absence of recombination hot-spots or cold-spots.

responsible primarily for differences in the epidemiology, host ranges, cell tropisms, transmission modes, replication levels and disease phenotypes displayed by BFDV and PiCV. It is also noteworthy that in *rep* 12/26 of these 'differentially favoured' codon sites are clustered within three genome regions (indicated by asterisks in Fig. 5); a pattern that may, for example, represent the locations of the main regions in Rep that are necessary for it to interact with and/or adapt to novel host species (Finsterbusch *et al.*, 2009).

Concluding remarks

In this study, we cloned and sequenced 30 PiCV isolates from a range of columbid birds sampled in Poland. Our analysis of these 30 isolates together with a further 17 available in public databases reveals that PiCV and BFDV likely have remarkably similar epidemiological characteristics and evolutionary dynamics. In addition to displaying similar degrees of diversity, isolates from both species display very low degrees of phylogenetic clustering based on sampling locations and host genotypes. There are also striking similarities between the nucleic acid secondary structures, signatures of natural selection and recombination patterns within the genomes of isolates from these two species.

As with many other ssDNA viruses it is likely that BFDV and PiCV have the capacity to evolve rapidly through both high mutation and high recombination rates (Duffy &

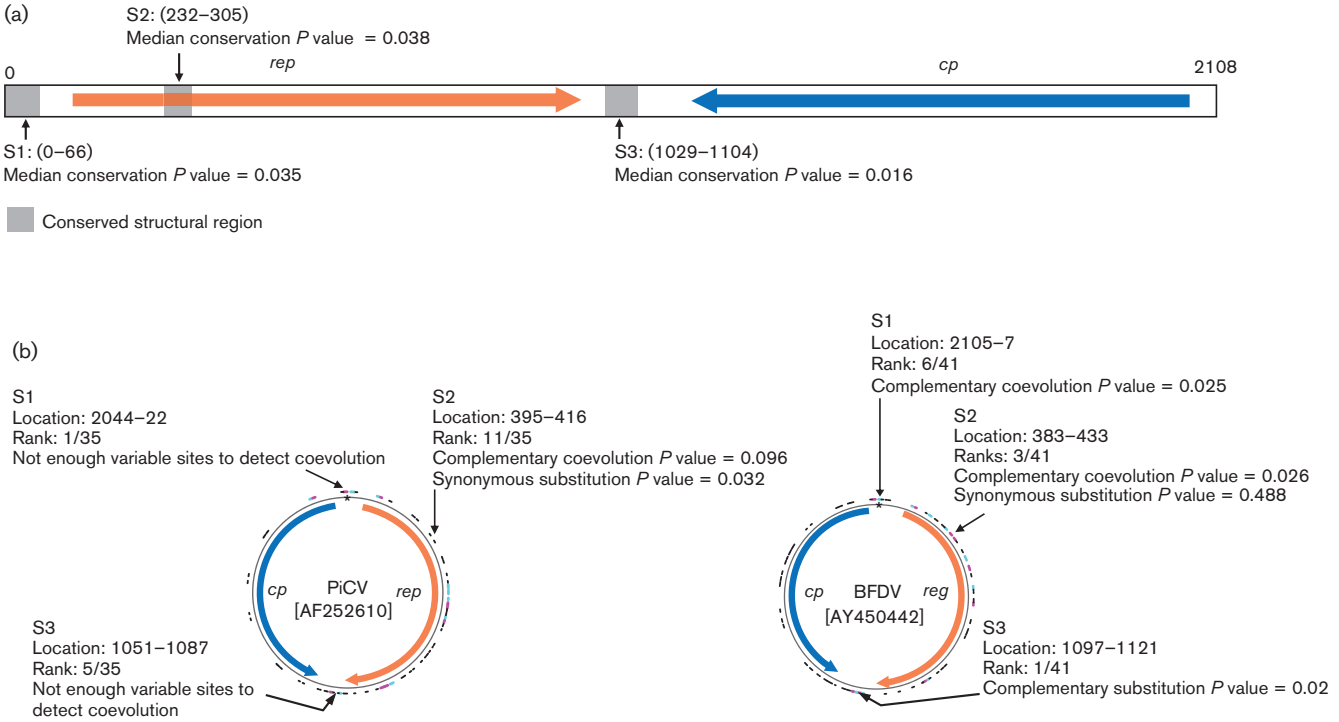

Fig. 4. Secondary structure conservation between PiCV and BFDV genomes. (a) Secondary structure conservation map of aligned PiCV and BFDV genomes, indicating three regions containing significantly conserved genomic secondary structural features (associated median conservation *P* values are given for each region). (b) Each genome is presented showing the location of its most conserved structural elements (from Muhire *et al.*, 2014), which were ranked based on base-pairing conservation scores, synonymous substitution rates and evidence of complementary coevolution favouring maintenance of base pairing. Structures are shown using arcs coloured in black or blue and pink (to distinguish the two complementary parts of the stem sequences for the 10 highest ranked structures). In the PiCV and BFDV datasets, S1 at the origin of replication ranks first and sixth, respectively, S2 within the *rep* gene ranks 11th and third, respectively, and S3 within the intergenic region between the *rep* and *cp* stop codons ranks fifth and first, respectively.

Holmes, 2008, 2009; Duffy *et al.*, 2008; Firth *et al.*, 2009; Grigoras *et al.*, 2010; Harkins *et al.*, 2009, 2014; Lefeuvre *et al.*, 2010). Given the capacity of viruses both to move rapidly across the world and infect a wide range of different host species/genotypes, it is perhaps surprising that there have been no major BFDV or PiCV global pandemics. Although such pandemics are unprecedented for bird-infecting circoviruses, they have occurred recently with the mammal-infecting circovirus, porcine circovirus type 2 (PCV-2). Despite PCV-2 being regarded initially as a ubiquitous but relatively harmless virus, a novel globally distributed pathogenic PCV-2 variant has emerged over the past 15–20 years to become one of the most important pathogens of pigs (Allan *et al.*, 2012; Grau-Roma *et al.*, 2011; Meng, 2012; Patterson & Opriessnig, 2010; Segalés *et al.*, 2013). It might therefore be prudent, wherever it is reasonably possible, to minimize the risks of pathogenic PiCV and BFDV variants evolving. Whilst nothing can be done to curb the rate at which BFDV and PiCV accumulate mutations during replication, it should be possible to reduce opportunities for these viruses to recombine by, for example, routinely testing for PiCV and isolating infected domestic birds, and establishing quarantine guidelines for inter-continental pigeon transportation.

METHODS

Sample collection and DNA extraction. Samples were collected from both living (n=177; from the cloaca) and dead birds (n=147; from either the liver, spleen or bursa of Fabricius) between 2002 and 2013. The 324 pigeons that were sampled included domestic carrier and fancy pigeons (n=264), and feral pigeons (n=60) from different regions of Poland. DNA was extracted using a Janus automated workstation (Perkin Elmer) and a NucleoMag Tissue kit (Macherey-Nagel) according to the manufacturer's instructions.

Detection of PiCV by PCR and recovery of complete genomes. Aliquots of 2 µl total extracted DNA were used for PCR-based screening using the HotStarTaq Plus Master Mix kit (Qiagen) with primers targeting a ~325 bp region of the capsid protein gene, PiCV2-s 5′-TTGAAAGGTTTTCAGCCTGGC-3′ and PiCV2-as 5′-AGGAGACGAAGGACACGCCTC-3′ (Freick *et al.*, 2008), using the following protocol: 95 °C for 5 min, then 30 cycles of 94 °C (60 s), 60 °C (60 s) and 72 °C (90 s), and final elongation after the last cycle 72 °C for 10 min.

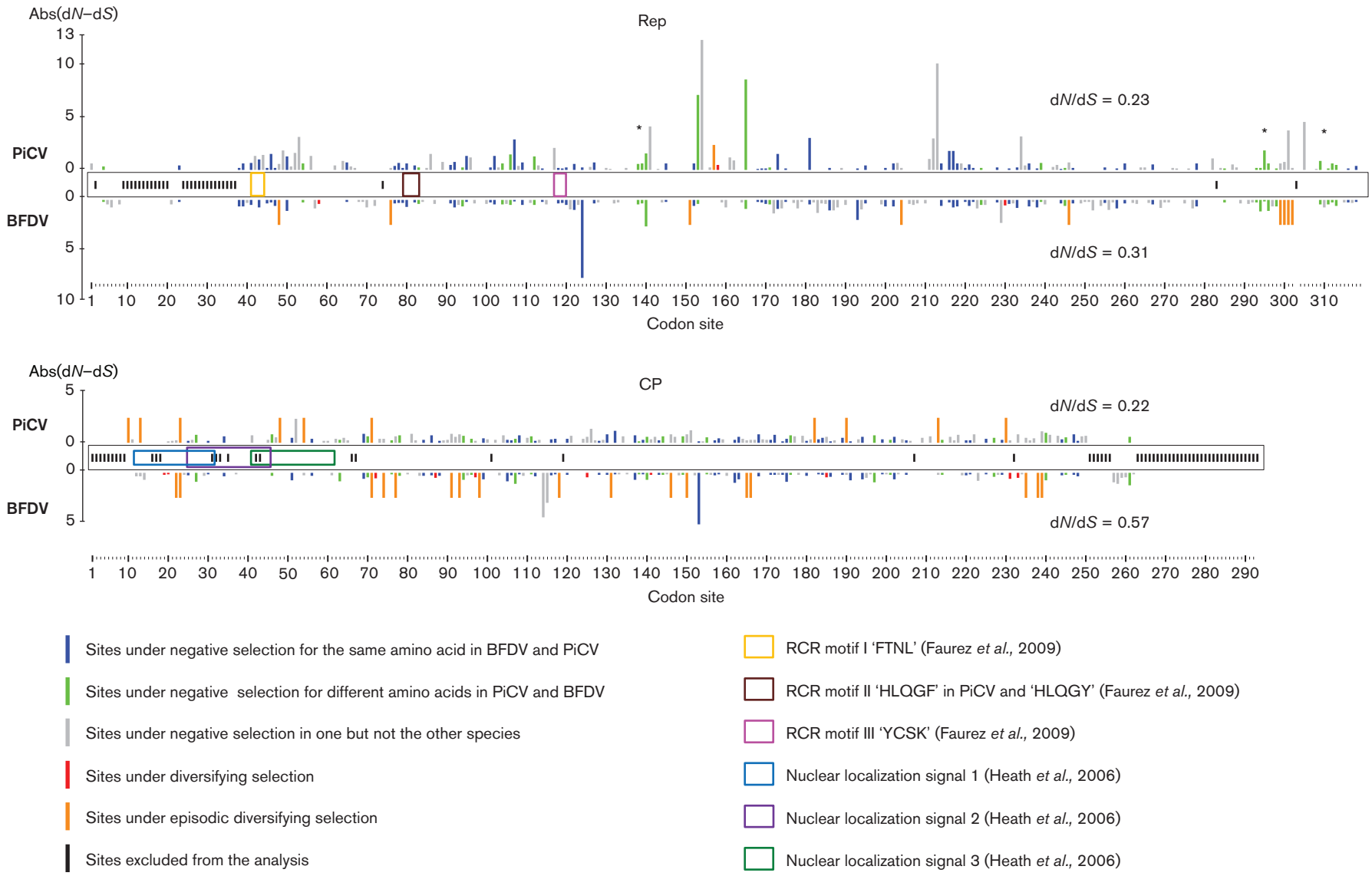

Fig. 5. Patterns of natural selection acting at Rep and CP codon sites in PiCV and BFDV. Rep and CP codon sites are shown aligned for PiCV and BFDV where for each site absolute (Abs) values of d*N*–d*S* are plotted (as determined by the FUBAR method). Significantly positive values are indicated in red (indicating positive selection), and significantly negative values are plotted in green and blue (indicating negative selection). Sites in orange are those identified by the MEME method as displaying evidence of episodic positive selection (i.e. sites identified by MEME but not FUBAR as evolving under positive selection). Overall d*N*/d*S* ratios <1 indicate that, as expected, all these genes are evolving under predominantly negative selection. The locations of codon sites are indicated based on the BFDV/PiCV codon alignments and the positions of known protein domains are indicated relative to these alignments. d*N*, Non-synonymous substitution rate; d*S*, synonymous substitution rate. Asterisks indicate clusters of codon sites under negative selection for different amino acids.

Total DNA (1 µl) of each PiCV-positive was enriched for circular DNA by rolling circle amplification (RCA) using TempliPhi (GE Healthcare) as described previously (Julian *et al.*, 2012, 2013; Massaro *et al.*, 2012). The concatenated RCA DNA was cut with *Xmn*I endonuclease and the resulting fragments (various sizes <2 kb) were cloned into the pJET1.2 plasmid vector (Thermo Fisher). The resulting clones were sequenced by primer walking at Macrogen (Korea).

All PiCV full genomes (n=17) available in public databases were downloaded from GenBank, and these were aligned with the RCA-amplified and cloned *Xmn*I fragments. Based on this alignment we designed abutting degenerate primers in the replication associated protein gene (back-to-back; PiCV-AV-F 5′-TCGCGCGAGASTTCA-GTGARAT-3′ and PiCV-AV-R 5′-CYTCSGYCATTGCTCTTCCGG-CTTTC-3′) capable of amplifying the complete genomes of all of the detected PiCV variants. KAPA HiFi HotStart DNA polymerase (Kapa Biosystems) was used with the primer pair PiCV-AV-F and PiCV-AV-R to amplify the complete genomes of the various detected PiCV variants using the following thermal cycling conditions: 94 °C for 2 min, then 25 cycles of 98 °C (20 s), 60 °C (20 s) and 72 °C (2 min), and a final extension of 72 °C for 2 min. The PCR amplicons were ligated to the pJET1.2 vector and the resulting plasmids, obtained for each of the PiCV genome isolates from a single transformed *E. coli* colony, were Sanger sequenced at Macrogen using pJET1.2-specific sequencing primers and internal sequencing primers.

Phylogenetic analysis. The complete PiCV genomes were assembled using DNA baser (version 3.5.2; Heracle BioSoft) and DNAMAN (version 5.2.9; Lynnon Biosoft). The assembled PiCV sequences (n=30) together with those available in public databases (n=17) were aligned using MUSCLE (Edgar, 2004) implemented in MEGA5 (Tamura *et al.*, 2011). Pairwise nucleotide identities (p-distances with pairwise deletion of gaps) were calculated using SDT v.1.0 (Muhire *et al.*, 2013). A neighbour-joining phylogenetic tree of the complete genomes was computed using MEGA5 with the Jukes–Cantor substitution model with 1000 bootstrap replicates. A maximum-likelihood phylogeny of the full genomes with recombinant regions removed was inferred using PHYML (Guindon *et al.*, 2010) with the GTR+G4 nucleotide substitution model (selected as the best fit model by jModelTest; Posada, 2009) and 1000 bootstrap replicates. Branches of the neighbour-joining and maximum-likelihood phylogenetic trees with <70 % bootstrap support were collapsed using Mesquite v2.75 (http://mesquiteproject.org/) and both the trees were rooted with a BFDV genome (GenBank accession number AF071878).

Recombination analysis. We analysed the 47 PiCV genomes for recombination using RDP4 (Martin *et al.*, 2010) with default settings using the recombination detection methods RDP (Martin & Rybicki, 2000), GENECONV (Padidam *et al.*, 1999), BOOTSCAN (Martin *et al.*, 2005), MAXCHI (Smith, 1992), CHIMAERA (Posada & Crandall, 2001), SISCAN (Gibbs *et al.*, 2000) and 3SEQ (Boni *et al.*, 2007). Recombination signals associated with P values <0.05 that were detected by three or more methods and which had associated phylogenetic support for recombination were accepted as credible evidence of recombination events. Sequences in the analysed dataset that most closely resembled the parental sequences of recombinants were defined as either 'minor parents' or 'major parents' based on the size of the genome fragments that these sequences had contributed to the detected recombinants (with the major parent contributing the larger fragment and the minor parent the smaller).

Predicting secondary structure conservation between PiCV and BFDV genomes. In total, 35 PiCV sequences and 35 BFDV sequences representative of the entire breadth of known PiCV and BFDV diversity were aligned using MAFFT (Katoh & Standley, 2013). The individual genomic secondary structures of each of these sequences were inferred computationally using the UNAfold program (Markham & Zuker, 2008) with the sequence conformation set as circular with a folding temperature of 37 °C (the lower bound of the natural physiological temperature of pigeons). The gap positions introduced by the MAFFT multiple sequence alignment procedure were maintained within each of the computationally inferred secondary structures, thus forming a multiple alignment of secondary structures. In order to determine which positions were conserved structurally within the alignment, all pairs of aligned secondary structures (each pair consisting of exactly one PiCV genomic secondary structure and exactly one BFDV genomic secondary structure) were taken and a substructure window of length 75 nt was moved along the pair of secondary structures with a mountain similarity metric (Moulton *et al.*, 2000) being calculated at each position. A permutation test was applied to each windowed position in each pair of aligned secondary structures. Specifically, a substructure conservation P value was obtained by computing the mountain similarity metric for each of 1000 permuted windows and then calculating the fraction of permuted similarity scores that were greater than or equal to the real similarity score. Permutations were generated by offsetting the coordinates of the original substructure by a random amount. As a result, each position within the secondary structure alignments had a corresponding list of conservation P values indicating the degree of structural conservation between all pairs of substructures of length 75 nt at that position. The overall conservation P value for all pairs of substructures at a given position within the entire alignment was calculated by taking the median conservation P value at that position, thus providing the degree of substructure conservation within each 75 nt window along the length of the PiCV+BFDV secondary structure alignment. The interpretation of the median P value at a position is as follows: a median P value of p indicates that at least 50 % of comparisons between each pair of sequences in the BFDV and PiCV multiple sequence alignments had a P value of p or lower. For example, if it was found that the median P value at a particular position was 0.05 this would imply that in at least 50 % of the $35 \times 35 = 1225$ BFDV versus PiCV pairwise comparisons (there were 35 sequences in each of the BFDV and PiCV datasets), a conservation P value of $\leqslant 0.05$ was detected at that position. A median P value of $\leqslant 0.05$ at a particular window position was considered significant evidence of a conserved substructure at that position. Adjacent windows with median P values of $\leqslant 0.05$ were combined and the median P value recalculated based on the combined window.

Detecting natural selection within the Rep and CP codon alignments. We detected and compared natural selection signals within the codon alignments of PiCV and BFDV *rep* and *cp* genes (produced using MEGA from ORFs extracted from all the available BFDV and PiCV full genome sequences) using two different codon-by-codon selection detection approaches. Although aligned together so as to identify homologous codon sites between the PiCV and BFDV genes, genes from the two species were analysed separately to identify recombination breakpoints (with GARD; Kosakovsky Pond *et al.*, 2006), and codon sites evolving under both positive and negative selection using FUBAR (Murrell *et al.*, 2013) and MEME (Murrell *et al.*, 2012). Whereas sites detectably evolving under positive selection by both FUBAR and MEME are likely to be evolving under consistent positive selection across all lineages, those sites which were only detected by MEME to be evolving under positive selection are likely only to be evolving under positive selection in particular lineages and not others.

ACKNOWLEDGEMENTS

A. V. and D. P. M. are supported by the National Research Foundation, South Africa. B. M. M. is funded by the University of Cape Town. T. S. is funded by the Statutory Research of Ministry of Science and Higher Education, Poland. The molecular work was

funded through a Polish Ministry of Education and Science grant (3 P06K 027 23) awarded to T. P. and a University of Cape Town (South Africa) Block Grant awarded to A. V.

REFERENCES

Abadie, J., Nguyen, F., Groizeleau, C., Amenna, N., Fernandez, B., Guereaud, C., Guigand, L., Robart, P., Lefebvre, B. & Wyers, M. (2001). Pigeon circovirus infection: pathological observations and suggested pathogenesis. *Avian Pathol* **30**, 149–158.

Allan, G., Krakowka, S., Ellis, J. & Charreyre, C. (2012). Discovery and evolving history of two genetically related but phenotypically different viruses, porcine circoviruses 1 and 2. *Virus Res* **164**, 4–9.

Bennett, S. N., Holmes, E. C., Chirivella, M., Rodriguez, D. M., Beltran, M., Vorndam, V., Gubler, D. J. & McMillan, W. O. (2006). Molecular evolution of dengue 2 virus in Puerto Rico: positive selection in the viral envelope accompanies clade reintroduction. *J Gen Virol* **87**, 885–893.

Biagini, P., Bendinelli, M., Hino, S., Kakkola, L., Mankertz, A., Niel, C., Okamoto, H., Raidal, S., Teo, C. G. & Todd, D. (2012). Family - Circoviridae. In *Virus Taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses*, pp. 343–349. Edited by A. M. Q. King, E. Lefkowitz, M. J. Adams & E. B. Carstens. San Diego, CA: Elsevier.

Blanchard, P., Mahé, D., Cariolet, R., Keranflec'h, A., Baudouard, M. A., Cordioli, P., Albina, E. & Jestin, A. (2003). Protection of swine against post-weaning multisystemic wasting syndrome (PMWS) by porcine circovirus type 2 (PCV2) proteins. *Vaccine* **21**, 4565–4575.

Boni, M. F., Posada, D. & Feldman, M. W. (2007). An exact nonparametric method for inferring mosaic structure in sequence triplets. *Genetics* **176**, 1035–1047.

Cheung, A. K. (2009). Homologous recombination within the capsid gene of porcine circovirus type 2 subgroup viruses via natural co-infection. *Arch Virol* **154**, 531–534.

Coletti, M., Franciosini, M. P., Asdrubali, G. & Passamonti, F. (2000). Atrophy of the primary lymphoid organs of meat pigeons in Italy associated with circoviruslike particles in the bursa of Fabricius. *Avian Dis* **44**, 454–459.

Cságola, A., Lorincz, M., Tombácz, K., Wladár, Z., Kovács, E. & Tuboly, T. (2012). Genetic diversity of pigeon circovirus in Hungary. *Virus Genes* **44**, 75–79.

Duchatel, J. P., Todd, D., Curry, A., Smyth, J. A., Bustin, J. C. & Vindevogel, H. (2005). New data on the transmission of pigeon circovirus. *Vet Rec* **157**, 413–415.

Duchatel, J. P., Todd, D., Smyth, J. A., Bustin, J. C. & Vindevogel, H. (2006). Observations on detection, excretion and transmission of pigeon circovirus in adult, young and embryonic pigeons. *Avian Pathol* **35**, 30–34.

Duffy, S. & Holmes, E. C. (2008). Phylogenetic evidence for rapid rates of molecular evolution in the single-stranded DNA begomovirus tomato yellow leaf curl virus. *J Virol* **82**, 957–965.

Duffy, S. & Holmes, E. C. (2009). Validation of high rates of nucleotide substitution in geminiviruses: phylogenetic evidence from East African cassava mosaic viruses. *J Gen Virol* **90**, 1539–1547.

Duffy, S., Shackelton, L. A. & Holmes, E. C. (2008). Rates of evolutionary change in viruses: patterns and determinants. *Nat Rev Genet* **9**, 267–276.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**, 1792–1797.

Esteban, D. J. & Hutchinson, A. P. (2011). Genes in the terminal regions of orthopoxvirus genomes experience adaptive molecular evolution. *BMC Genomics* **12**, 261.

Faurez, F., Dory, D., Grasland, B. & Jestin, A. (2009). Replication of porcine circoviruses. *Virol J* **6**, 60.

Finsterbusch, T., Steinfeldt, T., Doberstein, K., Rödner, C. & Mankertz, A. (2009). Interaction of the replication proteins and the capsid protein of porcine circovirus type 1 and 2 with host proteins. *Virology* **386**, 122–131.

Firth, C., Charleston, M. A., Duffy, S., Shapiro, B. & Holmes, E. C. (2009). Insights into the evolutionary history of an emerging livestock pathogen: porcine circovirus 2. *J Virol* **83**, 12813–12821.

Freick, M., Müller, H. & Raue, R. (2008). Rapid detection of pigeon herpesvirus, fowl adenovirus and pigeon circovirus in young racing pigeons by multiplex PCR. *J Virol Methods* **148**, 226–231.

Gibbs, M. J., Armstrong, J. S. & Gibbs, A. J. (2000). Sister-scanning: a Monte Carlo procedure for assessing signals in recombinant sequences. *Bioinformatics* **16**, 573–582.

Grau-Roma, L., Fraile, L. & Segalés, J. (2011). Recent advances in the epidemiology, diagnosis and control of diseases caused by porcine circovirus type 2. *Vet J* **187**, 23–32.

Grigoras, I., Timchenko, T., Grande-Pérez, A., Katul, L., Vetten, H. J. & Gronenborn, B. (2010). High variability and rapid evolution of a nanovirus. *J Virol* **84**, 9105–9117.

Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W. & Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* **59**, 307–321.

Harkins, G. W., Delport, W., Duffy, S., Wood, N., Monjane, A. L., Owor, B. E., Donaldson, L., Saumtally, S., Triton, G. & other authors (2009). Experimental evidence indicating that mastreviruses probably did not co-diverge with their hosts. *Virol J* **6**, 104.

Harkins, G. W., Martin, D. P., Christoffels, A. & Varsani, A. (2014). Towards inferring the global movement of beak and feather disease virus. *Virology* **450-451**, 24–33.

Heath, L., Martin, D. P., Warburton, L., Perrin, M., Horsfield, W., Kingsley, C., Rybicki, E. P. & Williamson, A. L. (2004). Evidence of unique genotypes of beak and feather disease virus in southern Africa. *J Virol* **78**, 9277–9284.

Heath, L., Williamson, A. L. & Rybicki, E. P. (2006). The capsid protein of beak and feather disease virus binds to the viral DNA and is responsible for transporting the replication-associated protein into the nucleus. *J Virol* **80**, 7219–7225.

Julian, L., Lorenzo, A., Chenuet, J. P., Bonzon, M., Marchal, C., Vignon, L., Collings, D. A., Walters, M., Jackson, B. & Varsani, A. (2012). Evidence of multiple introductions of beak and feather disease virus into the Pacific islands of Nouvelle-Caledonie (New Caledonia). *J Gen Virol* **93**, 2466–2472.

Julian, L., Piasecki, T., Chrzastek, K., Walters, M., Muhire, B., Harkins, G. W., Martin, D. P. & Varsani, A. (2013). Extensive recombination detected among beak and feather disease virus isolates from breeding facilities in Poland. *J Gen Virol* **94**, 1086–1095.

Katoh, K. & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* **30**, 772–780.

Kosakovsky Pond, S. L., Posada, D., Gravenor, M. B., Woelk, C. H. & Frost, S. D. (2006). GARD: a genetic algorithm for recombination detection. *Bioinformatics* **22**, 3096–3098.

Krapez, U., Slavec, B., Steyer, A. F., Pintaric, S., Dobeic, M., Rojs, O. Z. & Dovc, A. (2012). Prevalence of pigeon circovirus infections in feral pigeons in Ljubljana, Slovenia. *Avian Dis* **56**, 432–435.

Lefeuvre, P., Lett, J. M., Varsani, A. & Martin, D. P. (2009). Widely conserved recombination patterns among single-stranded DNA viruses. *J Virol* **83**, 2697–2707.

Lefeuvre, P., Martin, D. P., Harkins, G., Lemey, P., Gray, A. J., Meredith, S., Lakay, F., Monjane, A., Lett, J. M. & other authors (2010). The spread of tomato yellow leaf curl virus from the Middle East to the world. *PLoS Pathog* **6**, e1001164.

Mankertz, A., Hattermann, K., Ehlers, B. & Soike, D. (2000). Cloning and sequencing of columbid circovirus (CoCV), a new circovirus from pigeons. *Arch Virol* **145**, 2469–2479.

Markham, N. R. & Zuker, M. (2008). UNAFold: software for nucleic acid folding and hybridization. *Methods Mol Biol* **453**, 3–31.

Martin, D. & Rybicki, E. (2000). RDP: detection of recombination amongst aligned sequences. *Bioinformatics* **16**, 562–563.

Martin, D. P., Posada, D., Crandall, K. A. & Williamson, C. (2005). A modified bootscan algorithm for automated identification of recombinant sequences and recombination breakpoints. *AIDS Res Hum Retroviruses* **21**, 98–102.

Martin, D. P., Lemey, P., Lott, M., Moulton, V., Posada, D. & Lefeuvre, P. (2010). RDP3: a flexible and fast computer program for analyzing recombination. *Bioinformatics* **26**, 2462–2463.

Massaro, M., Ortiz-Catedral, L., Julian, L., Galbraith, J. A., Kurenbach, B., Kearvell, J., Kemp, J., van Hal, J., Elkington, S. & other authors (2012). Molecular characterisation of beak and feather disease virus (BFDV) in New Zealand and its implications for managing an infectious disease. *Arch Virol* **157**, 1651–1663.

Meng, X. J. (2012). Spread like a wildfire – the omnipresence of porcine circovirus type 2 (PCV2) and its ever-expanding association with diseases in pigs. *Virus Res* **164**, 1–3.

Moulton, V., Zuker, M., Steel, M., Pointon, R. & Penny, D. (2000). Metrics on RNA secondary structures. *J Comput Biol* **7**, 277–292.

Muhire, B., Martin, D. P., Brown, J. K., Navas-Castillo, J., Moriones, E., Zerbini, F. M., Rivera-Bustamante, R., Malathi, V. G., Briddon, R. W. & Varsani, A. (2013). A genome-wide pairwise-identity-based proposal for the classification of viruses in the genus Mastrevirus (family Geminiviridae). *Arch Virol* **158**, 1411–1424.

Muhire, B. M., Golden, M., Murrell, B., Lefeuvre, P., Lett, J. M., Gray, A., Poon, A. Y., Ngandu, N. K., Semegni, Y. & other authors (2014). Evidence of pervasive biologically functional secondary-structures within the genomes of eukaryotic single-stranded DNA viruses. *J Virol* **88**, 1972–1989.

Murrell, B., Wertheim, J. O., Moola, S., Weighill, T., Scheffler, K. & Kosakovsky Pond, S. L. (2012). Detecting individual sites subject to episodic diversifying selection. *PLoS Genet* **8**, e1002764.

Murrell, B., Moola, S., Mabona, A., Weighill, T., Sheward, D., Kosakovsky Pond, S. L. & Scheffler, K. (2013). FUBAR: a fast, unconstrained Bayesian approximation for inferring selection. *Mol Biol Evol* **30**, 1196–1205.

Padidam, M., Sawyer, S. & Fauquet, C. M. (1999). Possible emergence of new geminiviruses by frequent recombination. *Virology* **265**, 218–225.

Paré, J. A., Brash, M. L., Hunter, D. B. & Hampson, R. J. (1999). Observations on pigeon circovirus infection in Ontario. *Can Vet J* **40**, 659–662.

Patterson, A. R. & Opriessnig, T. (2010). Epidemiology and horizontal transmission of porcine circovirus type 2 (PCV2). *Anim Health Res Rev* **11**, 217–234.

Pogranichnyy, R. M., Yoon, K. J., Harms, P. A., Swenson, S. L., Zimmerman, J. J. & Sorden, S. D. (2000). Characterization of immune response of young pigs to porcine circovirus type 2 infection. *Viral Immunol* **13**, 143–153.

Posada, D. (2009). Selection of models of DNA evolution with jModelTest. *Methods Mol Biol* **537**, 93–112.

Posada, D. & Crandall, K. A. (2001). Evaluation of methods for detecting recombination from DNA sequences: computer simulations. *Proc Natl Acad Sci U S A* **98**, 13757–13762.

Rahaus, M., Desloges, N., Probst, S., Loebbert, B., Lanttermann, W. & Wolff, M. H. (2008). Detection of beak and feather disease virus DNA in embryonated eggs of psittacine birds. *Vet Med (Praha)* **53**, 53–58.

Raue, R., Schmidt, V., Freick, M., Reinhardt, B., Johne, R., Kamphausen, L., Kaleta, E. F., Müller, H. & Krautwald-Junghanns, M.-E. (2005). A disease complex associated with pigeon circovirus infection: young pigeon disease syndrome. *Avian Pathol* **34**, 418–425.

Rigby, M. A., Holmes, E. C., Pistello, M., Mackay, A., Leigh Brown, A. J. & Neil, J. C. (1993). Evolution of structural proteins of feline immunodeficiency virus: molecular epidemiology and evidence of selection for change. *J Gen Virol* **74**, 425–436.

Segalés, J., Kekarainen, T. & Cortey, M. (2013). The natural history of porcine circovirus type 2: from an inoffensive virus to a devastating swine disease? *Vet Microbiol* **165**, 13–20.

Seibert, S. A., Howell, C. Y., Hughes, M. K. & Hughes, A. L. (1995). Natural selection on the *gag*, *pol*, and *env* genes of human immunodeficiency virus 1 (HIV-1). *Mol Biol Evol* **12**, 803–813.

Smith, J. M. (1992). Analyzing the mosaic structure of genes. *J Mol Evol* **34**, 126–129.

Stenzel, T. A., Pestka, D., Tykałowski, B., Śmiałek, M. & Koncicki, A. (2012). Epidemiological investigation of selected pigeon viral infections in Poland. *Vet Rec* **171**, 562.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. & Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28**, 2731–2739.

Todd, D., Weston, J. H., Soike, D. & Smyth, J. A. (2001). Genome sequence determinations and analyses of novel circoviruses from goose and pigeon. *Virology* **286**, 354–362.

Todd, D., Fringuelli, E., Scott, A. N., Borghmans, B. J., Duchatel, J. P., Shivaprasad, H. L., Raidal, S. R., Abadie, J. X., Franciosini, M. P. & Smyth, J. A. (2008). Sequence comparison of pigeon circoviruses. *Res Vet Sci* **84**, 311–319.

Varsani, A., Regnard, G. L., Bragg, R., Hitzeroth, I. I. & Rybicki, E. P. (2011). Global genetic diversity and geographical and host-species distribution of beak and feather disease virus isolates. *J Gen Virol* **92**, 752–767.

Woods, L. W. & Latimer, K. S. (2000). Circovirus infection of nonpsittacine birds. *J Avian Med Surg* **14**, 154–163.

Woods, L. W., Latimer, K. S., Barr, B. C., Niagro, F. D., Campagnoli, R. P., Nordhausen, R. W. & Castro, A. E. (1993). Circovirus-like infection in a pigeon. *J Vet Diagn Invest* **5**, 609–612.

Woods, L. W., Latimer, K. S., Niagro, F. D., Riddell, C., Crowley, A. M., Anderson, M. L., Daft, B. M., Moore, J. D., Campagnoli, R. P. & Nordhausen, R. W. (1994). A retrospective study of circovirus infection in pigeons: nine cases (1986–1993). *J Vet Diagn Invest* **6**, 156–164.

Zhang, Z., Lu, C., Wang, Y., Wang, S., Dai, D., Chen, Z. & Fan, H. (2011). Molecular characterization and epidemiological investigation of Pigeon circovirus isolated in eastern China. *J Vet Diagn Invest* **23**, 665–672.